Alzheimer's & Dementia 12 (2016) 645-653

Featured Article

# Crowdsourced estimation of cognitive decline and resilience in Alzheimer's disease

Genevera I. Allen[a], Nicola Amoroso[b,c], Catalina Anghel[d], Venkat Balagurusamy[e], Christopher J. Bare[f], Derek Beaton[g], Roberto Bellotti[b,c], David A. Bennett[h], Kevin L. Boehme[i], Paul C. Boutros[d,j,k], Laura Caberlotto[l], Cristian Caloian[d], Frederick Campbell[a], Elias Chaibub Neto[f], Yu-Chuan Chang[m], Beibei Chen[n], Chien-Yu Chen[o], Ting-Ying Chien[p], Tim Clark[q,r], Sudeshna Das[q,r], Christos Davatzikos[s], Jieyao Deng[t,u], Donna Dillenberger[e], Richard J. B. Dobson[v,w,ccc], Qilin Dong[t,u], Jimit Doshi[s], Denise Duma[x], Rosangela Errico[y], Guray Erus[s], Evan Everett[a], David W. Fardo[z,aa], Stephen H. Friend[f], Holger Fröhlich[bb], Jessica Gan[a], Peter St George-Hyslop[cc], Satrajit S. Ghosh[dd,ee], Enrico Glaab[ff], Robert C. Green[gg], Yuanfang Guan[hh,ii,jj], Ming-Yi Hong[o], Chao Huang[kk], Jinseub Hwang[ll], Joseph Ibrahim[kk], Paolo Inglese[mm], Anandhi Iyappan[oo,bb], Qijia Jiang[a], Yuriko Katsumata[aa], John S. K. Kauwe[i,*], Arno Klein[f,**], Dehan Kong[kk], Roland Krause[ff], Emilie Lalonde[d], Mario Lauria[l], Eunjee Lee[kk], Xihui Lin[d], Zhandong Liu[a], Julie Livingstone[d], Benjamin A. Logsdon[f], Simon Lovestone[nn], Tsung-wei Ma[n], Ashutosh Malhotra[oo,bb], Lara M. Mangravite[f,**], Taylor J. Maxwell[pp], Emily Merrill[q], John Nagorski[a], Aishwarya Namasivayam[ff], Manjari Narayan[a], Mufassra Naz[oo,bb], Stephen J. Newhouse[v,qq], Thea C. Norman[f], Ramil N. Nurtdinov[rr], Yen-Jen Oyang[m], Yudi Pawitan[ss], Shengwen Peng[t,u], Mette A. Peters[f,**], Stephen R. Piccolo[i], Paurush Praveen[l,bb], Corrado Priami[l], Veronica Y. Sabelnykova[d], Philipp Senger[oo], Xia Shen[ss,aaa,bbb], Andrew Simmons[v], Aristeidis Sotiras[s], Gustavo Stolovitzky[tt,e], Sabina Tangaro[c], Andrea Tateo[b], Yi-An Tung[uu], Nicholas J. Tustison[vv], Erdem Varol[s], George Vradenburg[ww], Michael W. Weiner[xx], Guanghua Xiao[n], Lei Xie[yy], Yang Xie[n], Jia Xu[n], Hojin Yang[kk], Xiaowei Zhan[n], Yunyun Zhou[n], Fan Zhu[hh], Hongtu Zhu[kk], Shanfeng Zhu[t,u,zz], for the Alzheimer's Disease Neuroimaging Initiative

[a]*Department of Statistics and Electrical and Computer Engineering, Rice University, Houston, TX, USA*
[b]*Dipartimento di Fisica "M. Merlin", Università degli studi di Bari "A. Moro", Bari, Italy*
[c]*Sezione di Bari, Istituto Nazionale di Fisica Nucleare, Bari, Italy*
[d]*Ontario Institute for Cancer Research, Informatics and Bio-computing Program, MaRS Centre, Toronto, ON, Canada*
[e]*IBM Computational Biology Center, IBM Research, NY, USA*
[f]*Sage Bionetworks, Seattle, WA, USA*

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

*Corresponding author. Tel.: +1 801 422 2993; Fax: +1 801 422 0004.
**Lara M Mangravite: Tel.: +1 206 667 6044; Fax: +1 206 667 2062 Mette A. Peters: Tel.: +1 206 667 2113; Fax: +1 206 667 2062 Arno Klein: Tel.: + 1 917 512 5627
E-mail address: kauwe@byu.edu (J.S.K.K.), arno@binarybottle.com (A.K.), lara.mangravite@sagebase.org (L.M.M.), mette.peters@sagebase.org (M.A.P.)

http://dx.doi.org/10.1016/j.jalz.2016.02.006

[g]School of Behavioral and Brain Sciences, The University of Texas at Dallas, Richardson, TX, USA
[h]Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL, USA
[i]Department of Biology, Brigham Young University, Provo, UT, USA
[j]Department of Medical Biophysics, University of Toronto, Toronto, Canada
[k]Department of Pharmacology & Toxicology, University of Toronto, Toronto, Canada
[l]The Microsoft Research, University of Trento Centre for Computational and Systems Biology (COSBI), Rovereto, Italy
[m]Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei, Taiwan
[n]Quantitative Biomedical Research Center, The University of Texas Southwestern Medical Center, Dallas, TX, USA
[o]Department of Bio-Industrial Mechatronics Engineering, National Taiwan University, Taipei, Taiwan
[p]Innovation Center for Big Data and Digital Convergence, Yuan Ze University, Taoyuan, Taiwan
[q]Department of Neurology, Massachusetts General Hospital, Cambridge, MA, USA
[r]Department of Neurology, Harvard Medical School, Boston, MA, USA
[s]Center for Biomedical Image Computing and Analytics, University of Pennsylvania, Philadelphia, PA, USA
[t]School of Computer Science, Fudan University, Shanghai, Shanghai, China
[u]Shanghai Key Lab of Intelligent Information Processing, Fudan University, Shanghai, Shanghai, China
[v]NIHR Biomedical Research Centre for Mental Health, Kings College London, London, UK
[w]Institute of Psychiatry, Psychology and Neuroscience, MRC Social, Genetic and Developmental Psychiatry Centre, Kings College London, London, UK
[x]Department of Pediatrics-Neurology, Baylor College of Medicine, Houston, TX, USA
[y]Università degli Studi di Genova, Genova, Italy
[z]Sanders-Brown Center on Aging, University of Kentucky, Lexington, KY, USA
[aa]Department of Biostatistics, University of Kentucky, Lexington, KY, USA
[bb]Bonn-Aachen International Center for IT, University of Bonn, Bonn, Germany
[cc]Cambridge Institute for Medical Research, University of Cambridge and University of Toronto, Cambridge, CB2, UK
[dd]McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA, USA
[ee]Department of Otology and Laryngology, Harvard Medical School, Boston, MA, USA
[ff]Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg
[gg]Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Broad Institute and Harvard Medical School, Boston, MA, USA
[hh]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA
[ii]Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, USA
[jj]Department of Internal Medicine, University of Michigan, Ann Arbor, MI, USA
[kk]Department of Biostatistics, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[ll]Department of Computer science and Statistics, Daegu University, Gyeongsan-si, Gyeongsangbuk-do, Republic of Korea
[mm]Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London, UK
[nn]Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford, UK
[oo]Fraunhofer Institute for Algorithms and Scientific Computing (SCAI), Department for Bioinformatics, Schloss Birlinghoven, Sankt Augustin, Germany
[pp]Computational Biology Institute, The George Washington University, Ashburn, VA, USA
[qq]Department of Biostatistics, Kings College London, London, UK
[rr]Department of Neuroimmunology, Foundation Institut de Recerca, Hospital Universitari Vall d'Hebron, Barcelona, Spain
[ss]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
[tt]Genetics and Genomics Sciences Department, Icahn School of Medicine at Mount Sinai, New York, NY, USA
[uu]Genome and systems biology degree program, National Taiwan University, Taipei, Taiwan
[vv]Department of Radiology and Medical Imaging, The University of Virginia, Charlottesville, VA, USA
[ww]Global CEO Initiative on Alzheimer's disease, Washington, DC, USA
[xx]Radiology, Medicine, Psychiatry, and Neurology, UCSF, SFVAMC, San Francisco, CA, USA
[yy]Department of Computer Science, Hunter College, The City University of New York, New York, NY, USA
[zz]Centre for Computational Systems Biology, Fudan University, Shanghai, China
[aaa]Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK
[bbb]MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK
[ccc]Farr Institute of Health Informatics Research, UCL Institute of Health Informatics, University College London, London WC1E 6BT, UK

**Abstract** Identifying accurate biomarkers of cognitive decline is essential for advancing early diagnosis and prevention therapies in Alzheimer's disease. The Alzheimer's disease DREAM Challenge was designed as a computational crowdsourced project to benchmark the current state-of-the-art in predicting cognitive outcomes in Alzheimer's disease based on high dimensional, publicly available genetic and structural imaging data. This meta-analysis failed to identify a meaningful predictor developed from either data modality, suggesting that alternate approaches should be considered for prediction of cognitive performance.



*Keywords:* Azheimer's disease; Biomarkers; Crowdsource; Big data; Bioinformatics; Cognitive decline; Imaging; Genetics

## 1. Introduction

The Alzheimer's disease DREAM challenge (http://dx.doi.org/10.7303/syn2290704) was designed to provide an unbiased assessment of current capabilities for estimation of cognition and prediction of cognitive decline using genetic and imaging data from public data resources using a crowdsourced approach. The ability to predict rate of cognitive decline—both before and after diagnosis—is essential to effective trial design for the development of therapies for Alzheimer's disease (AD) prevention and treatment. Major collaborative efforts in the field are assessing the association of genetic loci with AD diagnosis and the application of structural imaging for development of early biomarkers of diagnosis, but the utility of these approaches to estimate cognition or predict cognitive decline is not well established. This project was designed under the advisement of a panel of experts in the field to evaluate whether these questions could be meaningfully addressed with current methods given existing public data sources. To ensure that these questions were tested across a broad spectrum of the latest analytical approaches, the study was designed as a crowdsourced, community-based challenge in which participants were invited to address one or more of the following three questions [1]: The prediction of cognitive decline over time based on genetic data [2]. The prediction of resilience to cognitive decline in individuals with elevated amyloid burden based on genetic data [3]. The estimation of cognitive state based on structural magnetic resonance (MR) imaging data.

## 2. Results

### 2.1. Study design and data harmonization

To ensure that predictors were detecting true biological variation rather than study-specific technical variation, this project required inclusion of data from multiple study sources. Although genetic and imaging data have been generated within many rich longitudinal cohorts across the field, the procurement and harmonization of these data sets were a nontrivial problem that required solutions to overcome political, ethical, and technical barriers. For example, the generation of whole genome sequencing data across multiple AD cohorts within the NIH-funded AD sequencing project has resulted in a powerful resource for genetic analysis in the field but longitudinal information on cognitive traits is not readily available in those data sets. Despite limitations on data accessibility, multiple relevant data sources were identified and used in this project including the Alzheimer's Disease Neuroimaging Initiative (ADNI) [1], the Rush Alzheimer's Disease Center Religious Orders Study [2], Memory and Aging Project (ROS/MAP) [3], and the European AddNeuroMed [4] study, which is part of InnoMed, a precursor to the innovative medicines initiative. Data selection and processing were performed based on data availability across these three data sets. As such, cognition was defined using mini mental state examination (MMSE) scores [5], genetic data were provided based on imputation across array-based genotype data, and structural MR imaging data were reprocessed in each cohort using a common processing pipeline. Genetic and imaging data were supplemented with a limited set of covariates including diagnosis, initial MMSE score, age at the initial examination, years of education, gender, and *APOE* haplotype. Participants were provided with data from ADNI to train algorithms over a 4-month period and to ensure that participation was not limited by access to compute resources, they were offered use of the IBM zEnterprise cloud to perform analyses. The challenge generated significant interest with 527 individuals from around the world registered to participate. A leaderboard displayed accuracy of submissions throughout the duration of the challenge: 1157 submissions were made for question 1, 478 submissions for question 2, and 434 submissions for question 3. Thirty-two teams submitted final results that were scored based on prediction and/or estimation of blinded outcomes within ROS/MAP for genetic predictions and AddNeuroMed for imaging-based estimations (Fig. 1).

### 2.2. Genetic prediction of cognitive decline

The first challenge question assessed the ability of current methods to predict change in cognitive examination performance based on genetic data. High prediction accuracy would signal the potential for noninvasive biomarkers of cognition to have a major clinical impact on early AD diagnosis and prevention. Previous efforts to develop predictors of change in cognitive function have not succeeded in providing robust and replicable models [6–8]. Genetic variation has been demonstrated to influence AD status: rare genetic mutations at several loci are implicated in familial forms of early-onset disease [9], whereas common variation contributes 33% to variance in sporadic AD, and 22 loci have been implicated by large-scale genetic association analyses [10,11]. However, with the exception of the *APOE* ε4 haplotype, there has been little success in transforming these genetic associations into meaningful clinical predictions of cognitive decline. For this purpose, participants were challenged to predict 2-year changes in MMSE scores based on genotypes imputed from SNP array data. Participants trained their algorithms with 767 ADNI samples, and the algorithms' predictions were evaluated on a test set of 1175 ROS/MAP samples with blinded outcome measures. The algorithm with the best predictive performance at the midpoint of the challenge did not contain any genetic features beyond *APOE* haplotype. As the goal of this question was to assess genetic contribution to prediction of cognitive decline, this top-ranked algorithm was openly shared across teams as an interim baseline on which to incorporate additional genetic predictors (http://dx.doi.org/10.7303/syn2838779). Eighteen teams submitted final predictions. Most methods performed significantly better than a permutation-based random model prediction (Fig. 2A). A cluster of six methods performed significantly better than the others (including the interim baseline model) but were

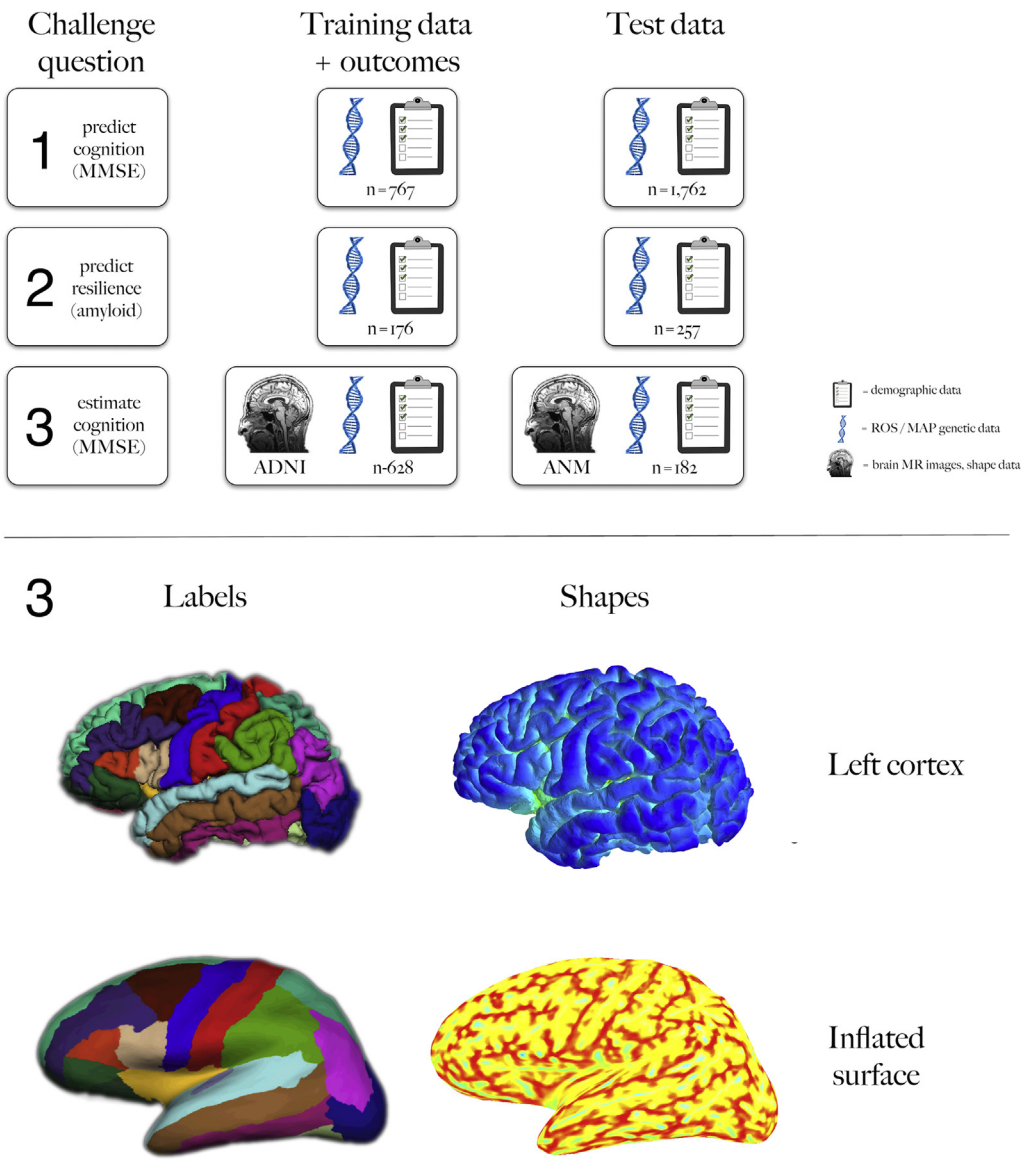


Fig. 1. Challenge overview. The top schematic summarizes the three challenge questions on the left column, the training data in the middle, and the test data on the right, including numbers of subjects. The symbols represent sources of data (demographic, ROS/MAP genetic, and ADNI or ANM brain images and shape information). The bottom panel provides example brain image labels and shape information derived from the Mindboggle software (http://mindboggle.info) provided to the participants for question 3. Anatomic labels for left cortical regions are shown on the left and just a couple of the cortical surface shape measures are shown on the right (travel depth on top and mean curvature below), for both uninflated and inflated surfaces (top and bottom rows, respectively).

statistically indistinguishable among themselves (Fig. 2D). Of these, the prediction with the best overall score (team GuanLab_umich from the University of Michigan) achieved a Pearson correlation of 0.382 and a Spearman correlation of 0.433 (the overall score was a rank-based combination of these two measures of performance; see online Supplement and Supplementary Methods: http://dx.doi.org/10.7303/syn3383106). However, no significant contribution of genetics beyond *APOE* haplotype to predictive performance was observed across any of the submissions. Given the small sample size, no conclusions can be inferred from this analysis regarding the existence of genetic loci associated with cognitive decline. Rather, these observations suggest that predictors of cognitive decline developed based on genetic data will not be useful within the clinical setting.

### 2.3. Genetic prediction of cognitive resilience

The second question challenged participants to identify genetic predictors that could distinguish individuals who exhibit resilience to AD pathology as defined by minimal change in cognitive function despite evidence of amyloid deposition [12,13]. Identification of genetic signatures predictive of cognitive resilience would aid in the elucidation of mechanisms that may confer resilience, providing a powerful tool to help advance AD prevention

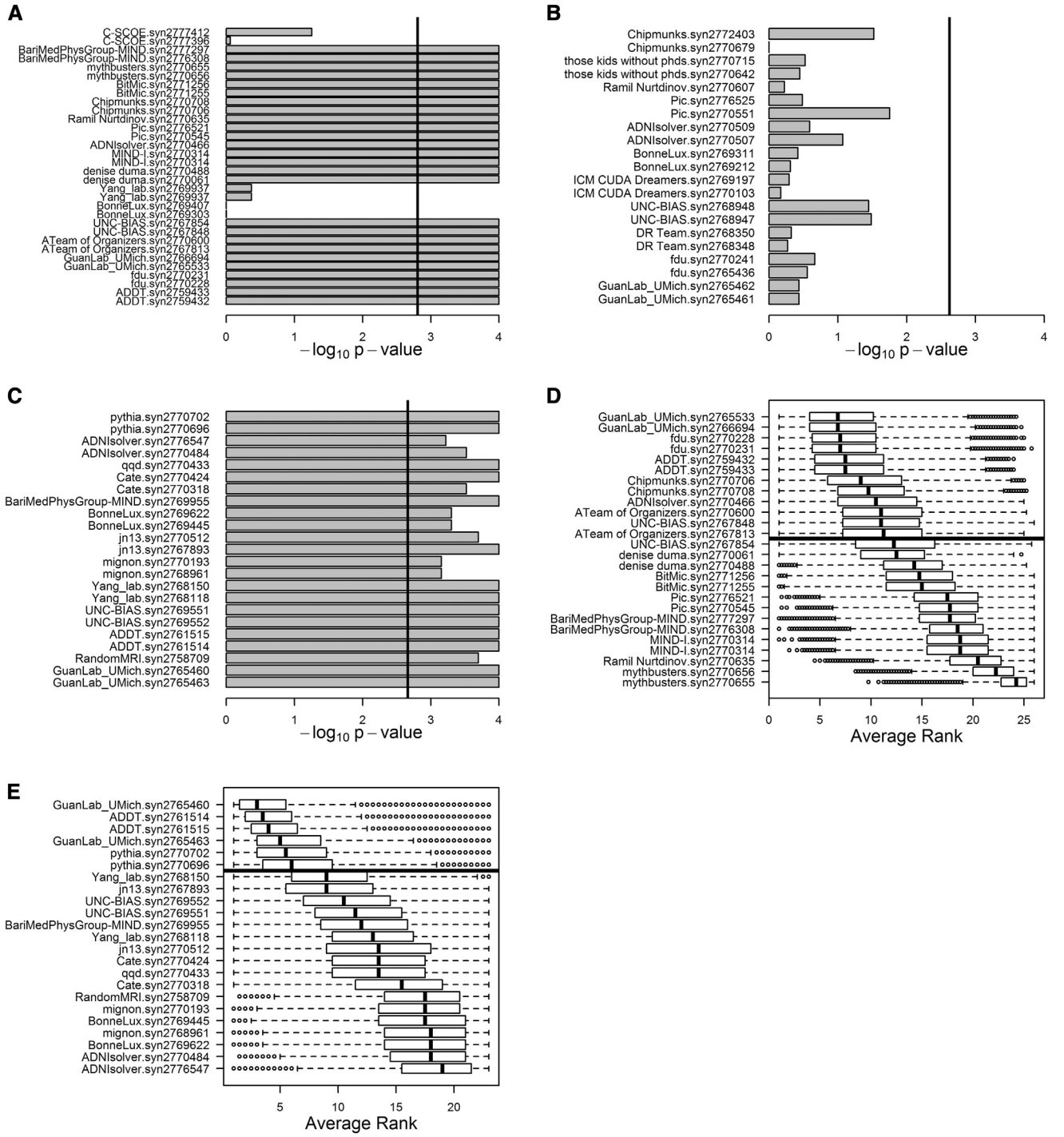


Fig. 2. Performance evaluation results. (A), (B), and (C) report the *P* values (in negative log 10 scale) for intersection union tests investigating which teams performed better than random for questions 1, 2, and 3, respectively. Explicitly, for question 1 (A), we tested the null hypothesis that at least one of the four correlation coefficients (namely Pearson/clinical, Pearson/clinical + genetics, Spearman/clinical, and Spearman/clinical + genetics) is equal to zero, against the alternative that all four correlation coefficients are larger than zero. Adopting a 0.05 significance level, 26 of the 32 submissions were statistically better than random, after Bonferroni multiple testing correction for 32 tests (submissions crossing the black vertical line). For question 2 (B), we tested the null hypothesis that balanced accuracy = 0.5 or AUC = 0.5, against the alternative that balanced accuracy > 0.5 and AUC > 0.5. In this case, no model performed significantly better than random, and, therefore, no best performer was declared. For question 3 (C), we tested the null hypothesis that Pearson's correlation (COR) or Lin's concordance correlation coefficient (CCC) are equal to zero, against the alternative that both COR and CCC are larger than zero. Adopting a 0.05 significance level, all 23 submissions were statistically better than random, after Bonferroni correction. For all three questions, the *P* values were computed from an empirical null distribution based on 10,000 permutations. (D) and (E) report the bootstrapped assessment of ranks for questions 1 and 3, respectively. Samples were resampled with replacement from the original data (true outcome and team's predictions), and the ranks of the different teams were reassessed in each of 100,000 resamplings. Submissions were sorted according to the median of their bootstrapped average ranking distributions. The black horizontal line represents the posterior odds cutoff from the Bayesian analysis. Teams above the black line are statistically tied to the top-ranked model, according to a posterior odds threshold of 3.

strategies and treatment development. Eleven teams submitted predictions of resilience based on a training set derived from 176 ADNI subjects. Evaluations were made using data derived from 257 individuals from the ROS/MAP data. Despite using the largest such public data set assembled to date, participants were unable to develop algorithms with predictive performances significantly better than random (see Fig. 2B, online Supplement and Supplementary Methods in Synapse: http://dx.doi.org/10.7303/syn3383106). Although it is likely that the study was underpowered due to small sample size and trait heterogeneity, this result suggests that information about cognitive resilience is not easily discoverable from SNP analysis.

### 2.4. Structural imaging-based estimation of cognition

The third question challenged participants to estimate cognitive state using structural brain image data (Fig. 1, lower panel). Brain imaging has emerged as a powerful method for monitoring neurodegeneration, and there is a great enthusiasm in the field to make use of images for diagnosis and prediction. There have been many attempts in the past to correlate changes in brain shape with disease progression and/or diagnosis, conventionally using measures of volume for a given brain region [14,15]. More detailed shape measures of image features including cortical thickness, curvature, and depth have also been found to be relevant to a variety of neurologic conditions [16]. Participants were challenged to estimate MMSE scores based on structural brain images, or shape information derived from these images. Participants trained algorithms using ADNI data (N = 628) and were evaluated using AddNeuromed data (N = 182) for which they were blind to outcome measures. To engage as many participants as possible from both within and beyond the neuroimaging community, the data were provided both as raw MR images and as tables containing shape measures (volume, thickness, area, curvature, depth, and so forth) for every labeled brain region. Thirteen teams submitted estimates for final evaluation, and all teams performed better than a random model (see online Supplement and Supplementary Methods in Synapse: http://dx.doi.org/10.7303/syn3383106). Three teams performed significantly better than the others (teams GuanLab_umich from the University of Michigan, ADDT from the Karolinska Institute and Pythia from the University of Pennsylvania; Fig. 2C) but were statistically indistinguishable from one another and tied for top average rank (Fig. 2E). The algorithm that generated the best absolute mean combined rank (Team GuanLab_umich) achieved a concordance correlation coefficient of 0.569 and Pearson's correlation of 0.573 (the overall score was a rank-based combination of these two measures of performance). The most common features that contributed heavily to the MMSE estimates across the algorithms were hippocampal volume and entorhinal thickness, corroborating prior work [17–19]. The top three teams also found that inclusion of shape measures of the entorhinal cortex (volume, curvature, surface area, travel, and geodesic depth) improved overall estimation. Other features that contributed to predictions within the top three teams' results included volume of inferior lateral ventricle and amygdala (see online Supplement and Supplementary Methods in Synapse: http://dx.doi.org/10.7303/syn3383106). These results validate an established relationship between structural imaging data and cognition. However, the correlative performance of these estimators was low suggesting that their application in the clinical setting may not be sufficient to inform patient care.

## 3. Discussion

The AD DREAM challenge provided a formalized assessment of the ability to develop meaningful predictions of cognitive performance from public genetic or imaging data using contemporary state-of-the-art predictive algorithms. Predictive performance across all three of the questions was modest, and most methods performed roughly equivalently. Given this uniform performance, we do not expect that the presented results are a failure of current modeling methods. A more likely explanation is that the data used to address these questions were inadequate to support these tasks. We also note that most research teams that participated in this challenge did not have expertise in the field of AD. Although the few teams that did possess this knowledge did not do better than the others, there remains the possibility that performance would have been improved by the inclusion of more domain experts.

### 3.1. Use of genetic information for cognitive prediction

The modest performance observed in the 3 questions focused on genetic analysis demonstrated that contemporary algorithms were not able to leverage genetic signal to make useful predictions for cognition. These results support the prevailing expectation that genetic variants of moderate to high frequency will not support viable biomarker development in AD [9–11]. Although heritability estimates and linkage studies have demonstrated that there is a considerable estimated genetic contribution to AD onset and progression [11,20,21], evidence both within the AD field and across other complex disease [22] traits has indicated that this overall genetic contribution is the aggregated compilation of a large number of loci with small—independent or epistatic—effects. Historically, this type of signal is difficult to capture in predictive models and unlikely to be useful in a diagnostic setting [23]. Furthermore, cognition is highly influenced by a host of nongenetic factors relating to lifestyle choices and accumulated exposures that were not represented across all these data sets and, in fact, are not fully captured in most cohorts [24–27]. Nongenetic contributions to cognitive performance may themselves provide an important base for successful predictions.

Within the context of genetic analysis, the absence of these factors from models confounds the ability to detect real genetic signal and impacts the ability to accurately model state-specific genetic contributions. As such, future inquiry into the use of genetic testing for prediction of cognitive performance and AD risk assessment may be better served by focusing on the contribution of rare genetic variation. Recently discovered disease-associated rare variants have larger effect sizes than common variants and confer 2- to 5-fold greater risk or protection in carriers relative to the general population [28–30]. Ongoing large-scale sequencing analyses will identify additional associated rare risk variants. In sufficient numbers, the aggregate prevalence would support the development of a genetic diagnostic containing a library of rare variants.

### 3.2. Use of structural imaging data for cognitive estimation

Although the inexpensive and noninvasive nature of genetic testing make this approach amenable to population-level disease screening, the resource-intensive nature of image-based testing is better positioned for careful evaluation of high-risk individuals. As such, these approaches are needed to provide a higher confidence estimate of cognitive performance. Although a variety of methods developed within the context of this challenge were able to successfully estimate cognition, none of these methods were sufficiently accurate to merit clinical consideration. These observations support previous work in the field [17,19] and highlight the imperfect relationship between brain structure and function. Newer imaging modalities that focus on brain function and/or pathology—such as FDG-PET [31] or tau imaging [32]—may prove more successful for assessing cognitive dysfunction.

### 3.3. Effective performance of meta-analysis across diverse cohorts

A major consideration for any meta-analysis is the issue of appropriate harmonization of data across disparate sources. Despite leveraging several of the most deeply phenotyped cohorts in the field, this challenge limited analysis to those traits that were in common across cohorts. Although this approach to data harmonization is standard practice for meta-analyses [10], it greatly reduced the depth of the information available for modeling and influenced the selection of cognitive measures for use as prediction outcomes. Because each cohort had performed a battery of study-specific tests, this greatly limited the ability for finer grained assessment across cognitive processes. A more sensible approach for future analyses may be to focus effort on more sophisticated methods to calibrate disparate cognitive phenotypes across cohorts [33]. Another undesirable consequence of the focus on traits measured in common was the inability to incorporate into model development the full spectrum of nongenetic and nonimaging factors that are known to influence cognitive performance [24–27]. This suggests the need for development of different approaches for integrating heterogeneous data and/or assessing replication across cohorts. Alternatively, smaller scale analyses that prioritize phenotypic depth over sample size may afford a more refined view of disease.

In summary, this challenge demonstrated that predictions of cognitive performance developed from genetic or structural imaging data were modest across a diverse set of contemporary modeling methods. Future efforts to identify clinically relevant predictors of cognition will benefit from a focus on alternative data sources and methods that work to incorporate greater phenotypic complexity.

## Acknowledgments

Author contributions: C.J.B., E.C.N., D.W.F., S.H.F., S.S.G., A.K., J.S.K.K., Y.K., B.A.L., L.M.M., T.J.M., T.C.N., M.A.P., G.S., G.V., and N.J.T. contributed to the challenge organization. V.B., D.D., P.S.H., and R.C.G. contributed with compute resources and scientific advice. G.I.A., N.A., C.A., D.B., R.B., K.L.B., P.C.B., L.C., C.C., F.C., Y.C.C., B.C., C.Y.C., T.Y.C., T.C., S.D., C.D., J.D., Q.D., J.D., D.D., R.E., G.E., E.E., H.F., J.G., E.G., Y.G., M.Y.H., C.H., J.H., J.I., P.I., A.I., Q.J., D.K., R.K., E.L., M.L., E.L., X.L., Z.L., J.L., T-w.M., A.M., E.M., J.N., A.N., M.N., M.N., R.N.N., Y.J.O., Y.P., S.P., S.R.P., P.P., C.P., V.Y.S., P.S., X.S., A.S., S.T., A.T., Y.A.T., E.V., G.X., L.X., Y.X., J.X., H.Y., X.Z., Y.Z., F.Z., H.Z., and S.Z. participated in the challenge community phase. D.A.B., R.J.B.D., S.L., S.J.N., A.S., and M.W.W. contributed with data used in the challenge.

The following people provided final submissions but did not participate in the community phase of the challenge: Lorna Barron (GIDAS, miRcore, 2929 Plymouth Rd. Suite 207, Ann Arbor, MI, USA), Oliver Barron(GIDAS, miRcore, 2929 Plymouth Rd. Suite 207, Ann Arbor, MI, USA), Riccardo Bellazzi (Electrical, Computer and Biomedical Engineering Department, Via Ferrata, 1, Pavia, Italy), Jungwoo Chang (GIDAS, miRcore, 2929 Plymouth Rd. Suite 207, Ann Arbor, MI, USA), Marianne H Cowherd (GIDAS, miRcore, 2929 Plymouth Rd. Suite 207, Ann Arbor, MI, USA), Grace Ganzel (GIDAS, miRcore, 2929 Plymouth Rd. Suite 207, Ann Arbor, MI, USA), Łukasz Grad (Interdisciplinary Centre for Mathematical and Computational Modelling, Pawińskiego 5A, Warsaw, Poland), Inhan Lee (GIDAS, miRcore, 2929 Plymouth Rd. Suite 207, Ann Arbor, MI, USA), Ivan Limongelli (Electrical, Computer and Biomedical Engineering Department, Via Ferrata, 1, Pavia, Italy), Simone Marini (Electrical, Computer and Biomedical Engineering Department, Via Ferrata, 1, Pavia, Italy), Szymon Migacz (Interdisciplinary Centre for Mathematical and Computational Modelling, Pawińskiego 5A, Warsaw, Poland), Ettore Rizzo (Electrical, Computer and Biomedical Engineering Department, Via Ferrata, 1, Pavia, Italy),

Witold R Rudnicki (Interdisciplinary Centre for Mathematical and Computational Modelling, Pawińskiego 5A, Warsaw, Poland; Department of Bioinformatics, University of Białystok, Ciołkowskiego 1M, Białystok, Poland), Andrzej Sułecki (Interdisciplinary Centre for Mathematical and Computational Modelling, Pawińskiego 5A, Warsaw, Poland), Leo Tunkle (GIDAS, miRcore, 2929 Plymouth Rd. Suite 207, Ann Arbor, MI, USA), Francesca Vitali (Electrical, Computer and Biomedical Engineering Department, Via Ferrata, 1, Pavia, Italy)

This study was supported by the following individuals and organizations: Alan Evans (McGill University), Gaurav Pandey (MSSM), Gil Rabinovici (UCSF), Kaj Blennow (Göteborg University), Kristine Yaffe (UCSF), Maria Isaac (EMA), Nolan Nichols (University of Washington), Paul Thompson (UCLA), Reisa Sperling (Harvard), Scott Small (Columbia), Guy Eakin (BrightFocus Foundation), Maria Carillo (Alzheimer's Association), Neil Buckholz (NIA), Alzheimer's Research UK, European Medicines Agency, Global CEO Initiative on Alzheimer's Disease, Pfizer, Inc, Ray and Dagmar Dolby Family Fund, Rosenberg Alzheimer's Project, Sanofi S.A, and Takeda Pharmaceutical Company Ltd, USAgainstAlzheimer's.

Study data were provided by the following groups: (1) The Alzheimer's Disease Neuroimaging Initiative (ADNI)—ADNI is funded by the National Institutes of Health (U01 AG024904), the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering and through generous contributions from the following: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. This research was also supported by NIH grants P30 AG010129 and K01 AG030514. (2) The Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago—Data collection was supported through funding by NIA grants P30AG10161, R01AG15819, R01AG17917, R01AG30146, R01AG36836, U01AG32984, and U01AG46152, the Illinois Department of Public Health, and the Translational Genomics Research Institute. (3) European AddNeuroMed study—The AddNeuroMed data are from a public-private partnership supported by EFPIA companies, SMEs and the EU under the FP6 programme. Clinical leads responsible for data collection are Iwona Kłoszewska (Lodz), Simon Lovestone (London), Patrizia Mecocci (Perugia), Hilkka Soininen (Kuopio), Magda Tsolaki (Thessaloniki), and Bruno Vellas (Toulouse).

## RESEARCH IN CONTEXT

1. Systematic review: Extensive literature searches using PubMed establish this as the largest study to date using demographic, clinical, imaging, and genetic data to predict cognitive decline and the first major instance of crowdsourcing analysis in AD.
2. Interpretation: Over 500 scientists worldwide in the analytical portion of the challenge, demonstrating the viability of crowdsourced approaches in AD research. Unfortunately, we were unable to detect meaningful predictors of either cognitive decline or resilience through this effort.
3. Future directions: This experiment in crowdsourcing AD analyses is an invaluable first-of-its-kind contribution that provides a snapshot of both the strengths and limitations in big data analytics in AD research. The relative inaccessibility and heterogeneity across data sources severely limits formalized integration. Mandates on data sharing, considerations of standardized data collection, and mechanisms to integrate heterogeneous data are necessary to address these issues. We anticipate that this work will initiate those discussions across the community.

## References

[1] Mueller SG, Weiner MW, Thal LJ, Petersen RC, Jack C, Jagust W, et al. The Alzheimer's disease neuroimaging initiative. Neuroimaging clinics of North America 2005;15:869–77. xi-xii.

[2] Bennett DA, Schneider JA, Arvanitakis Z, Wilson RS. Overview and findings from the religious orders study. Current Alzheimer research 2012;9:628–45.

[3] Bennett DA, Schneider JA, Buchman AS, Barnes LL, Boyle PA, Wilson RS. Overview and findings from the rush Memory and Aging Project. Current Alzheimer research 2012;9:646–63.

[4] Lovestone S, Francis P, Kloszewska I, Mecocci P, Simmons A, Soininen H, et al. AddNeuroMed–the European collaboration for the discovery of novel biomarkers for Alzheimer's disease. Annals of the New York Academy of Sciences 2009;1180:36–46.

[5] Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. Journal of psychiatric research 1975;12:189–98.

[6] Ercoli LM, Siddarth P, Dunkin JJ, Bramen J, Small GW. MMSE items predict cognitive decline in persons with genetic risk for Alzheimer's disease. Journal of geriatric psychiatry and neurology 2003;16:67–73.
[7] Hsiung GY, Alipour S, Jacova C, Grand J, Gauthier S, Black SE, et al. Transition from cognitively impaired not demented to Alzheimer's disease: an analysis of changes in functional abilities in a dementia clinic cohort. Dementia and geriatric cognitive disorders 2008; 25:483–90.
[8] Vemuri P, Wiste HJ, Weigand SD, Shaw LM, Trojanowski JQ, Weiner MW, et al. MRI and CSF biomarkers in normal, MCI, and AD subjects: predicting future clinical change. Neurology 2009; 73:294–301.
[9] Ridge PG, Ebbert MT, Kauwe JS. Genetics of Alzheimer's disease. BioMed research international 2013;2013:254954.
[10] Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nature genetics 2013;45:1452–8.
[11] Ridge PG, Mukherjee S, Crane PK, Kauwe JS, Alzheimer's Disease Genetics Consortium. Alzheimer's disease: analyzing the missing heritability. PloS one 2013;8:e79771.
[12] Bennett DA, Schneider JA, Arvanitakis Z, Kelly JF, Aggarwal NT, Shah RC, et al. Neuropathology of older persons without cognitive impairment from two community-based studies. Neurology 2006; 66:1837–44.
[13] Price JL, Morris JC. Tangles and plaques in nondemented aging and "preclinical" Alzheimer's disease. Annals of neurology 1999; 45:358–68.
[14] Davatzikos C, Xu F, An Y, Fan Y, Resnick SM. Longitudinal progression of Alzheimer's-like patterns of atrophy in normal older adults: the SPARE-AD index. Brain : a journal of neurology 2009;132(Pt 8):2026–35.
[15] Misra C, Fan Y, Davatzikos C. Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: results from ADNI. NeuroImage 2009; 44:1415–22.
[16] Im K, Lee JM, Seo SW, Hyung Kim S, Kim SI, Na DL. Sulcal morphology changes and their relationship with cortical thickness and gyral white matter volume in mild cognitive impairment and Alzheimer's disease. NeuroImage 2008;43:103–13.
[17] Haight TJ, Jagust WJ, Alzheimer's Disease Neuroimaging Initiative. Relative contributions of biomarkers in Alzheimer's disease. Annals of epidemiology 2012 Dec;22:868–75.
[18] Nho K, Risacher SL, Crane PK, DeCarli C, Glymour MM, Habeck C, et al. Voxel and surface-based topography of memory and executive deficits in mild cognitive impairment and Alzheimer's disease. Brain imaging and behavior 2012;6:551–67.
[19] Thung KH, Wee CY, Yap PT, Shen D, Alzheimer's Disease Neuroimaging Initiative. Neurodegenerative disease diagnosis using incomplete multi-modality data via matrix shrinkage and completion. NeuroImage 2014;91:386–400.
[20] Escott-Price V, Sims R, Bannister C, Harold D, Vronskaya M, Majounie E, et al. Common polygenic variation enhances risk prediction for Alzheimer's disease. Brain : a journal of neurology 2015; 138(Pt 12):3673–84.
[21] Lee SH, Harold D, Nyholt DR, ANZGene Consortium, International Endogene Consortium, Genetic and Environmental Risk for Alzheimer's disease Consortium, et al. Estimation and partitioning of polygenic variation captured by common SNPs for Alzheimer's disease, multiple sclerosis and endometriosis. Hum Mol Genet 2013; 22:832–41.
[22] Chatterjee N, Wheeler B, Sampson J, Hartge P, Chanock SJ, Park JH. Projecting the performance of risk prediction based on polygenic analyses of genome-wide association studies. Nature genetics 2013; 45:400–5. 5e1–3.
[23] Manolio TA. Bringing genome-wide association findings into clinical use. Nat Rev Genet 2013;14:549–58.
[24] Scarmeas N, Stern Y, Mayeux R, Luchsinger JA. Mediterranean diet, Alzheimer disease, and vascular mediation. Arch Neurol 2006; 63:1709–17.
[25] Podewils LJ, Guallar E, Kuller LH, Fried LP, Lopez OL, Carlson M, et al. Physical activity, APOE genotype, and dementia risk: findings from the Cardiovascular Health Cognition Study. Am J Epidemiol 2005;161:639–51.
[26] Lindsay J, Laurin D, Verreault R, Hebert R, Helliwell B, Hill GB, et al. Risk factors for Alzheimer's disease: a prospective analysis from the Canadian Study of Health and Aging. Am J Epidemiol 2002; 156:445–53.
[27] Wang HX, Karp A, Winblad B, Fratiglioni L. Late-life engagement in social and leisure activities is associated with a decreased risk of dementia: a longitudinal study from the Kungsholmen project. Am J Epidemiol 2002;155:1081–7.
[28] Guerreiro R, Wojtas A, Bras J, Carrasquillo M, Rogaeva E, Majounie E, et al. TREM2 variants in Alzheimer's disease. N Engl J Med 2013;368:117–27.
[29] Jonsson T, Stefansson H, Steinberg S, Jonsdottir I, Jonsson PV, Snaedal J, et al. Variant of TREM2 associated with the risk of Alzheimer's disease. N Engl J Med 2013;368:107–16.
[30] Jonsson T, Atwal JK, Steinberg S, Snaedal J, Jonsson PV, Bjornsson S, et al. A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. Nature 2012;488:96–9.
[31] Gray KR, Wolz R, Heckemann RA, Aljabar P, Hammers A, Rueckert D, et al. Multi-region analysis of longitudinal FDG-PET for the classification of Alzheimer's disease. NeuroImage 2012 Mar; 60:221–9.
[32] James OG, Doraiswamy PM, Borges-Neto S. PET Imaging of Tau Pathology in Alzheimer's Disease and Tauopathies. Front Neurol 2015; 6:38.
[33] Gross AL, Sherva R, Mukherjee S, Newhouse S, Kauwe JS, Munsie LM, et al. Calibrating longitudinal cognition in Alzheimer's disease across diverse test batteries and datasets. Neuroepidemiology 2014;43:194–205.